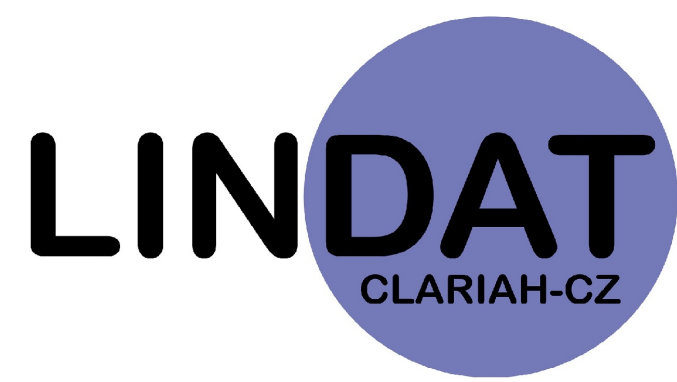




CHARLES
UNIVERSITY



CLARIN-DSpace

Ondřej Košarko



ÚSTAV FORMÁLNÍ
A APLIKOVANÉ LINGVISTIKY



matfyz

What is DSpace

- DSpace is the software of choice for academic, non-profit, and commercial organizations building open digital repositories.
- DSpace preserves and enables easy and open access to all types of digital content including text, images, moving images, mpegs and data sets.
- If you are building a literature repository, DSpace can be used “out of the box”
- If you are building a data repository in the context of language resources and CLARIN -> CLARIN-DSpace

DSpace versions recent history (Jun 7 2021)

- First version circa 2002 (jdk 1.4 timeframe, java was ~6 yrs old)
- DSpace 5.0 released Jan 16 2015
 - 5.10 released Oct 22 2018
 - Under security support until v8.0 is released
- DSpace 6.0 released Oct 24 2016
 - 6.3 released Jun 27 2018
 - Under bug-fix support until v7.0 is released. Under security support until v9.0 is released
- DSpace 7.0 release date Aug 2 2021 (or sooner)
 - Announced this week
- <https://wiki.lyrasis.org/display/DSPACE/Releases>

CLARIN-DSpace

- Currently build on DSpace 5.x
 - Started on 1.6/1.8.2 ~ 2010/2012
 - Uses XMLUI (see further slides) in the UI layer
 - Used by 10+ CLARIN centers
- Out of the box, preset, as a data repository.
 - Namely linguistic data & tools
 - preset to fulfill B-centre requirements
 - There's still a lot of paper pushing after you've installed the software
- Additions/enhancements
 - Licensing framework, customized submission workflow + CMDI, OpenAIRE (data repository) export (in oai-pmh), enhanced administration, configurable handles, localization enhancements, large files downloads, many small files zipped download, ...
 - A lot of those are interactive, ie. involve the UI layer and it's not just "a theme"

What's new in DSpace 6

- Major Java API refactor, supporting UUIDs and Hibernate
 - Most sites will not notice this major API refactor, as the upgrade is seamless. However, if you've performed major (Java-level) customizations, you may need to refactor your own customizations to use this newly refactored API
- Different configuration system
 - **Many** configurations now automatically reload
- File storage plugins, healthcheck, new features in xmlui, new features in rest api, updated OAI-PMH crosswalks, etc.
- Removing legacy browse, removing legacy search
- <https://wiki.lyrasis.org/display/DSPACE/DSpace+Release+6.0+Status>

What will be in DSpace 7.0

- New user interface based on angular, replacing jspui, xmlui (xmluiv2)
 - The goal of the new UI is to implement all **major features** of both XMLUI and JSPUI in a single, modern UI.
- Refactored REST API
- Redesigned submissions and workflow
- New configurable entities object model
- <https://wiki.lyrasis.org/display/DSPACE/DSpace+Release+7.0+Status>

What won't be in DSpace 7.0

- There's a rough roadmap for post 7.0 release (7.1->7.5) listing some 50 features
- ~12 new, eg.
 - Signposting
 - Resourcesync
 - Backup + restore of entities
- ~36 reimplementing features present in 6 which didn't make it into 7.0 (and were not dropped), eg.
 - Preview an item during workflow approval
 - Support for type-based submission forms (<type-bind> configuration)
 - List of Recently Added Items (Recent Submissions) on homepage

Difficulties upgrading CLARIN-DSpace

- The data & metadata are safe and can be exported. It's the processes (workflows etc.) that'll make the upgrade (to newer DSpace) difficult.
- 5.x is supported (security) until v8 is released
- 7.0, from my point of view, is not feature complete
 - But we can make the leap and implement the missing features, if there's enough man power
- 6.x is more of a technical release (no must have features)
 - It has the old UIs & the old REST api, build anything new around that and expect to reimplement that for next release
 - Needs to be tested performance wise, few mention degraded performance compared to 5.x